Amendments to the Specification

Please change the word selecs in newly numbered paragraph 0014 to selects as follows:

[0014]   In a third embodiment, system for analyzing a document is provided. The system receives a document that includes data and a document type. The document type has an associated physical structure. The system determines a logical structure of the document based in part on the data and [selecs] <u>selects</u> a subset of the data based on at least one of the group including the associated physical structure and the logical structure. The system also stores a document segment that includes the selected subset of the data.

Please add a period to the end of newly numbered paragraph 0028 as follows:

[0028]   The extraction tool 120 extracts relevant files from the various project repositories 110a, 110b, and 110x. In one embodiment, each repository has an associated repository type. The repository type defines the structure of the repository, such as the underlying directory structure for the repository. Additionally, the repository may be a simple repository consisting of a single directory, or a complex repository that may store metadata associated with each file kept in the repository. In one embodiment, the extraction tool 120 connects to each of the repositories 110a, 110b, and 110x through repository adapters 121a, 121b, and 121x. An adapter acts as an Application Programming Interface, or API, to the repository. For complex repositories, the repository adapter may allow for the extraction of metadata associated with the document.